

# A factorial moment distance and an application to the matching problem

G. Afendras\* and N. Papadatos†

Departments of Biostatistics and Biomedical Informatics, SUNY at Buffalo, 3435 Main Street, 726 Kimball Hall, School of Public Health and Health Professions, and School of Medicine, Buffalo, NY 14260

and

Department of Mathematics, University of Athens,  
Panepistemiopolis, 157 84 Athens, Greece.

## Abstract

In this note we introduce the notion of factorial moment distance for non-negative integer-valued random variables and we compare it with the total variation distance. Furthermore, we study the rate of convergence in the classical matching problem and in a generalized matching distribution.

*MSC:* Primary 60E05, 60E15; Secondary 44A10, 41A25.

*Key words and phrases:* Factorial moment distance; Total variation distance; Matching problem.

## 1 Introduction

Let  $\pi_n = (\pi_n(1), \dots, \pi_n(n))$  be a random permutation of  $T_n = \{1, 2, \dots, n\}$ , in the sense that  $\pi_n$  is uniformly distributed over the  $n!$  permutations of  $T_n$ . A number  $j$  is a fixed point of  $\pi_n$  if  $\pi_n(j) = j$ . Denote by  $Z_n$  the total number of fixed points of  $\pi_n$ ,

$$Z_n = \sum_{j=1}^n \mathbb{1}\{\pi_n(j) = j\},$$

where  $\mathbb{1}$  stands for the indicator function. The study of  $Z_n$  corresponds to the famous *matching problem*, introduced by Montmort in 1708. Obviously,  $Z_n$  can take the values  $0, 1, \dots, n-2, n$ , and its exact distribution, using standard combinatorial arguments, is found to be

$$\mathbb{P}(Z_n = j) = \frac{1}{j!} \sum_{k=0}^{n-j} \frac{(-1)^k}{k!}, \quad j = 0, 1, \dots, n-2, n.$$

---

\*Corresponding author. e-mail address: gafendra@buffalo.edu

†Work partially supported by the University of Athens' Research fund under Grant 70/4/5637.

It is obvious that  $Z_n$  converges in law to  $Z$ , where  $Z$  is the standard Poisson distribution,  $\text{Poi}(1)$ . Furthermore, the Poisson approximation is very accurate even for small  $n$  (evidence of this may be found in [Barbour et al., 1992](#)). Bounds on the error of the Poisson approximation in the matching problem, especially concerning the total variation distance, are also well-known. Recall that the total variation distance of any two rv's  $X_1$  and  $X_2$  is defined as

$$d_{\text{TV}}(X_1, X_2) = \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}(X_1 \in A) - \mathbb{P}(X_2 \in A)|,$$

where  $\mathcal{B}(\mathbb{R})$  is the Borel  $\sigma$ -algebra of  $\mathbb{R}$ . An appealing result is given by [Diaconis \(1987\)](#), who proved that  $d_{\text{TV}}(Z_n, Z) \leq \frac{2^n}{n!}$ . This bound has been improved by [DasGupta \(1999, 2005\)](#):

$$d_{\text{TV}}(Z_n, Z) \leq \frac{2^n}{(n+1)!}. \quad (1.1)$$

It can be seen that  $d_{\text{TV}}(Z_n, Z) \sim \frac{2^n}{(n+1)!}$ , where  $a_n \sim b_n$  means that  $\lim_n \frac{a_n}{b_n} = 1$ ; for a proof of a more general result see Theorem 3.2, below. Therefore, the bound (1.1) is of the correct order.

Consider now the sets of discrete rv's

$$\mathcal{D}_n := \{X : \mathbb{P}(X \in \{0, 1, \dots, n\}) = 1\}, \quad \mathcal{D}_\infty := \{X : \mathbb{P}(X \in \{0, 1, \dots\}) = 1\}.$$

Since the first  $n$  moments of  $Z_n$  and  $Z$  are identical and  $Z_n \in \mathcal{D}_n$ ,  $Z \in \mathcal{D}_\infty$ , one might think that

$$\inf_{X \in \mathcal{D}_n} \{d_{\text{TV}}(X, Z)\} \sim d_{\text{TV}}(Z_n, Z) \sim \frac{2^n}{(n+1)!}. \quad (1.2)$$

However, (1.2) is not true. In fact,

$$\min_{X \in \mathcal{D}_n} \{d_{\text{TV}}(X, Z)\} = 1 - e^{-1} \sum_{j=0}^n \frac{1}{j!} \sim \frac{e^{-1}}{(n+1)!}. \quad (1.3)$$

Indeed, for any  $X_1, X_2 \in \mathcal{D}_\infty$  with probability mass functions (pmf's)  $p_1$  and  $p_2$ , the total variation distance can be expressed as

$$d_{\text{TV}}(X_1, X_2) = \frac{1}{2} \sum_{j=0}^{\infty} |p_1(j) - p_2(j)| = \sum_{j=0}^{\infty} (p_1(j) - p_2(j))^+, \quad (1.4)$$

where  $x^+ = \max\{x, 0\}$ . Thus, for any  $X_1 \in \mathcal{D}_n$  (so that  $p_1(j) = 0$  for all  $j > n$ ), we get

$$d_{\text{TV}}(X_1, X_2) = \sum_{j=0}^n (p_1(j) - p_2(j))^+ \geq \sum_{j=0}^n (p_1(j) - p_2(j)) = 1 - \sum_{j=0}^n p_2(j) = \mathbb{P}(X_2 > n),$$

with equality if and only if  $p_1(j) \geq p_2(j)$ ,  $j = 0, 1, \dots, n$ . Applying the preceding inequality to  $p_2(j) = \mathbb{P}(Z = j) = \frac{e^{-1}}{j!}$  we get the equality in (1.3), and the minimum is attained by

any rv  $X \in \mathcal{D}_n$  with  $\mathbb{P}(X = j) \geq \frac{e^{-1}}{j!}$ ,  $j = 0, 1, \dots, n$ . Furthermore, the well-known Cauchy remainder in the Taylor expansion reads as

$$f(x) - \sum_{j=0}^n \frac{f^{(j)}(0)}{j!} x^j = \frac{1}{n!} \int_0^x (x-y)^n f^{(n+1)}(y) dy. \quad (1.5)$$

Applying (1.5) to  $f(x) = e^x$  we get the expression

$$1 - e^{-1} \sum_{j=0}^n \frac{1}{j!} = e^{-1} \left( e - \sum_{j=0}^n \frac{1}{j!} \right) = \frac{e^{-1}}{n!} \int_0^1 (1-y)^n e^y dy,$$

and by the obvious inequalities  $1 < e^y < 1 + (e-1)y$ ,  $0 < y < 1$ , we have

$$\frac{1}{n+1} < \int_0^1 (1-y)^n e^y dy < \frac{1}{n+1} \left( 1 + \frac{e-1}{n+2} \right).$$

It follows that

$$\frac{e^{-1}}{(n+1)!} < \min_{X \in \mathcal{D}_n} \{d_{\text{TV}}(X, Z)\} = 1 - e^{-1} \sum_{j=0}^n \frac{1}{j!} < \frac{e^{-1}}{(n+1)!} \left( 1 + \frac{e-1}{n+2} \right),$$

and therefore,  $\min_{X \in \mathcal{D}_n} \{d_{\text{TV}}(X, Z)\} \sim \frac{e^{-1}}{(n+1)!}$ .

In the present note we introduce and study a class of factorial moment distances,  $\{d_\alpha, \alpha > 0\}$ . These metrics are designed to capture the discrepancy among discrete distributions with finite moment generating function in a neighborhood of zero and, in addition, they satisfy the desirable property  $\min_{X \in \mathcal{D}_n} \{d_\alpha(X, Z)\} = d_\alpha(Z_n, Z)$ . In Section 3 we study the rate of convergence in a generalized matching problem, and we present closed form expansions and sharp inequalities for the factorial moment distance and the variational distance.

## 2 The factorial moment distance

We start with the following observation: For the rv's  $Z$  and  $Z_n$ ,

$$\mathbb{E}(Z)_k = 1 \quad \text{and} \quad \mathbb{E}(Z_n)_k = \mathbb{1}_{\{k \leq n\}}, \quad k = 0, 1, \dots, \quad (2.1)$$

where  $\mathbb{E}(X)_k$  denotes the  $k$ -th order descending factorial moment of  $X$  [for each  $x \in \mathbb{R}$ ,  $(x)_0 = 1$  and  $(x)_k = x(x-1) \cdots (x-k+1)$ ,  $k = 1, 2, \dots$ ]. For a proof of a more general result see Lemma 3.1, below.

The factorial moment distance will be defined in a suitable sub-class of discrete random variables, as follows. For each  $t \geq 0$  we define

$$\mathcal{X}(t) := \{X \in \mathcal{D}_\infty : \text{there exists } t' > t \text{ such that } P_X(1+t') < \infty\}, \quad (2.2)$$

where  $P_X(u) = \mathbb{E} u^X$  is the probability generating function of  $X$ . Also, we define

$$\mathcal{X}(\infty) := \bigcap_{t \in [0, \infty)} \mathcal{X}(t) = \{X \in \mathcal{D}_\infty : P_X(1 + t') < \infty \text{ for any } t' > 0\}. \quad (2.3)$$

Note that if  $X \in \mathcal{D}_n$  for some  $n$  then  $X \in \mathcal{X}(t)$  for each  $t \in [0, \infty]$ ; therefore, each  $\mathcal{X}(t)$  is non-empty. For  $0 \leq t_1 < t_2 \leq \infty$  it is obvious that  $\mathcal{X}(t_2) \subset \mathcal{X}(t_1)$ ; that is, the family  $\{\mathcal{X}(t), 0 \leq t \leq \infty\}$  is decreasing in  $t$ .

If  $X \in \mathcal{X}(0)$  then there exists a  $t' > 0$  such that  $P_X(1 + t') < \infty$ , i.e.,  $\mathbb{E} e^{\theta X} < \infty$  where  $\theta = \ln(1 + t') > 0$ . Since  $X$  is non-negative,  $\mathbb{E} e^{\theta X} < \infty$  implies that  $\mathbb{E} e^{uX} < \infty$  for all  $u \in (-\theta, \theta)$ , which means that  $X$  has finite moment generating function at a neighborhood of zero. Therefore,  $X$  has finite moments of any order and its pmf is characterized by its moments; equivalently,  $X$  has finite descending factorial moment of any order and its pmf is characterized by these moments. This enables the following

**Definition 2.1.** (a) Let  $X_1, X_2 \in \mathcal{X}(0)$ . For  $\alpha > 0$  we define the *factorial moment distance of order  $\alpha$*  of  $X_1, X_2$  by

$$d_\alpha(X_1, X_2) := \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!} |\mathbb{E}(X_1)_k - \mathbb{E}(X_2)_k|. \quad (2.4)$$

(b) Let  $X \in \mathcal{X}(0)$  and  $\{X_n\}_{n=1}^\infty \subset \mathcal{X}(0)$ . We say that  $X_n$  *converges in factorial moment distance of order  $\alpha$*  to  $X$ , in symbols  $X_n \rightarrow_\alpha X$ , if

$$d_\alpha(X_n, X) \rightarrow 0, \text{ as } n \rightarrow \infty.$$

One can easily check that the function  $d_\alpha : \mathcal{X}(0) \times \mathcal{X}(0) \rightarrow [0, \infty]$  is a distance. Obviously,  $X_n \rightarrow_\alpha X$  implies that the moments of  $X_n$  converge to the corresponding moments of  $X$ . Since every  $X \in \mathcal{X}(0)$  is characterized by its moments, it follows that the  $d_\alpha$  convergence (for any  $\alpha > 0$ ) is stronger than the convergence in law; the later is equivalent to the convergence in total variation – see Wang (1991). Of course, the converse is not true even in  $\mathcal{X}(\infty)$ . For example, consider the rv  $X$  with  $\mathbb{P}(X = 0) = 1$ , and the sequence of rv's  $\{X_n\}_{n=1}^\infty$ , where each  $X_n$  has pmf

$$p_n(j) = \begin{cases} 1 - \frac{1}{n} & , \quad j = 0, \\ \frac{1}{n} & , \quad j = n. \end{cases}$$

It is obvious that  $\{X, X_1, X_2, \dots\} \subset \mathcal{X}(\infty)$ , and the total variation distance is

$$d_{\text{tv}}(X_n, X) = \frac{1}{n} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Moreover, since  $\mathbb{E}(X)_k = 0$  and  $\mathbb{E}(X_n)_k = (n-1)_{k-1} \mathbb{1}\{k \leq n\}$  for all  $k = 1, 2, \dots$ , the  $d_\alpha$  distance does not converge to zero:

$$d_\alpha(X_n, X) = \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k!} (n-1)_{k-1} \mathbb{1}\{k \leq n\} > \frac{\alpha}{2} (n-1) \mathbb{1}\{2 \leq n\} \rightarrow \infty, \quad \text{as } n \rightarrow \infty.$$

**Remark 2.1.** Let  $X \in \mathcal{D}_n \setminus \{Z_n\}$ . It is obvious that  $\mathbb{E}(X)_k = 0$  for all  $k > n$ , and we can find an index  $k \in \{1, 2, \dots, n\}$  such that  $\mathbb{E}(X)_k \neq 1$ . From (2.1) and (2.4) we see that  $d_\alpha(X, Z) > d_\alpha(Z_n, Z)$ . Hence,

$$\inf_{X \in \mathcal{D}_n} \{d_\alpha(X, Z)\} = d_\alpha(Z_n, Z), \quad \text{for all } \alpha > 0.$$

**Proposition 2.1.** Let  $0 < \alpha_1 < \alpha_2$  and  $X_1, X_2 \in \mathcal{X}(0)$ .

(a)  $d_{\alpha_1}(X_1, X_2) \leq d_{\alpha_2}(X_1, X_2)$ .

(b) We cannot find a constant  $C = C(\alpha_1, \alpha_1) < 1$  such that for all  $X_1, X_2 \in \mathcal{X}(0)$ ,  $d_{\alpha_1}(X_1, X_2) \leq C d_{\alpha_2}(X_1, X_2)$ .

*Proof.* (a) is obvious. To see (b), it suffices to consider  $X_1$  and  $X_2$  with  $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_2 = 1) = 1$ . Then  $d_\alpha(X_1, X_2) = 1$  for every  $\alpha > 0$ .  $\square$

From (a) of the preceding proposition,  $X_n \rightarrow_{\alpha_2} X$  implies  $X_n \rightarrow_{\alpha_1} X$  for every  $\alpha_1 < \alpha_2$ . In the sequel we shall show that for any  $\alpha \geq 2$ , the inequality  $d_{\text{tv}}(X_n, X) \leq d_\alpha(X_n, X)$  holds true, provided  $\{X, X_1, X_2, \dots\} \subseteq \mathcal{X}(1)$ . To this end, we shall make use of the following “moment inversion” formula.

**Lemma 2.1.** If  $X \in \mathcal{X}(1)$  then its pmf  $p$  can be written as

$$p(j) = \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!} \binom{k}{j} \mathbb{E}(X)_k, \quad j = 0, 1, \dots \quad (2.5)$$

*Proof.* By the assumption  $X \in \mathcal{X}(1)$ , we can find a number  $t' > 1$  such that  $\mathbb{E}(1 + t')^X = \sum_{j=0}^{\infty} (1 + t')^j p(j) < \infty$ . Since  $X$  is non-negative, its probability generating function admits a Taylor expansion around 0 with radius of convergence  $R \geq 1 + t' > 2$ , i.e.,  $P(u) = \sum_{j=0}^{\infty} u^j p(j) \in \mathbb{R}$ ,  $|u| < R$ . It is well known that  $\frac{d^k}{du^k} P(u) \Big|_{u=1} = \mathbb{E}(X)_k$ , and since  $P$  admits a Taylor expansion around 1 with radius of convergence  $R' \geq t' > 1$ , we have

$$P(u) = \sum_{k=0}^{\infty} \frac{\mathbb{E}(X)_k}{k!} (u-1)^k, \quad |u-1| < R'.$$

Using the preceding expansion and the fact that  $0 \in (1 - R', 1 + R')$  we get

$$p(j) = \frac{1}{j!} \cdot \frac{d^j}{du^j} P(u) \Big|_{u=0} = \frac{1}{j!} \sum_{k=j}^{\infty} \frac{(u-1)^{k-j}}{(k-j)!} \mathbb{E}(X)_k \Big|_{u=0} = \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{j!(k-j)!} \mathbb{E}(X)_k,$$

completing the proof.  $\square$

It should be noted that the condition  $X \in \mathcal{X}(t)$  for some  $t \in [0, 1)$  is not sufficient for (2.5). As an example, consider the geometric rv  $X$  with pmf  $p(j) = 2^{-j-1}$ ,  $j = 0, 1, \dots$ . It is clear that  $X \notin \mathcal{X}(1)$ , but  $X \in \mathcal{X}(t)$  for each  $t \in [0, 1)$ . The factorial moments of  $X$  are  $\mathbb{E}(X)_k = k!$ ,  $k = 0, 1, \dots$ , and the rhs of (2.5),  $\sum_{k=j}^{\infty} (-1)^{k-j} \binom{k}{j}$ , is a non-convergent series.

**Theorem 2.1.** If  $X_1, X_2 \in \mathcal{X}(1)$  then  $d_{\text{tv}}(X_1, X_2) \leq d_\alpha(X_1, X_2)$  for each  $\alpha \geq 2$ .

*Proof.* In view of Proposition 2.1(a), it is enough to prove the desired result for  $\alpha = 2$ . By (1.4) and (2.5) we get

$$d_{\text{tv}}(X_1, X_2) = \frac{1}{2} \sum_{j=0}^{\infty} \left| \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!} \binom{k}{j} (\mathbb{E}(X_1)_k - \mathbb{E}(X_2)_k) \right| \leq \frac{1}{2} \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} \frac{1}{k!} \binom{k}{j} |\mathbb{E}(X_1)_k - \mathbb{E}(X_2)_k|.$$

Interchanging the order of summation according to Tonelli's Theorem, we have

$$d_{\text{tv}}(X_1, X_2) \leq \frac{1}{2} \sum_{k=0}^{\infty} \sum_{j=0}^k \frac{1}{k!} \binom{k}{j} |\mathbb{E}(X_1)_k - \mathbb{E}(X_2)_k| = \sum_{k=0}^{\infty} \frac{2^{k-1}}{k!} |\mathbb{E}(X_1)_k - \mathbb{E}(X_2)_k|.$$

The proof is completed by the fact that  $\mathbb{E}(X_1)_0 = \mathbb{E}(X_2)_0 = 1$ .  $\square$

Theorem 2.1 quantifies the fact that for any  $\alpha \geq 2$ , the  $d_\alpha$  convergence (in  $\mathcal{X}(1)$ ) implies the convergence in total variation, and provides convenient bounds for the rate of the total variation convergence. However, we note that such convenient bounds do not hold for  $\alpha < 2$ . In fact, for given  $\alpha \in (0, 2)$  and  $t \geq 0$ , we cannot find a finite constant  $C = C(\alpha, t) > 0$  such that  $d_{\text{tv}}(X_1, X_2) \leq C d_\alpha(X_1, X_2)$  for all  $X_1, X_2 \in \mathcal{X}(t)$ ; see Remark 3.1, below.

### 3 An application to a generalized matching problem

Consider the classical matching problem where, now, we record only a proportion of the matches, due to a random censoring mechanism. The censoring mechanism decides independently to every individual match. Specifically, when a particular match occurs, the mechanism counts this match with probability  $\lambda$ , independently of the other matches, and ignores this match with probability  $1 - \lambda$ , where  $0 < \lambda \leq 1$ . We are now interested on the number  $Z_n(\lambda)$  of the counted matches. The case  $\lambda = 1$  corresponds to the classical matching problem where all coincidences are recorded, so that  $Z_n = Z_n(1)$ .

The probabilistic formulation is as follows: Let  $\pi_n = (\pi_n(1), \dots, \pi_n(n))$  be a random permutation of  $\{1, \dots, n\}$ , as in the Introduction. Let also  $J_1(\lambda), \dots, J_n(\lambda)$  be iid Bernoulli( $\lambda$ ) rv's, independent of  $\pi_n$ . The number  $Z_n(\lambda)$  of the recorded coincidences can be written as

$$Z_n(\lambda) = \sum_{i=1}^n J_i(\lambda) \mathbb{1}\{\pi_n(i) = i\}.$$

Let  $A_i = \{J_i(\lambda) = 1\}$ ,  $B_i = \{\pi_n(i) = i\}$ ,  $E_i = A_i \cap B_i$ ,  $i = 1, \dots, n$ . Then  $Z_n(\lambda)$  presents the number of the events  $E$ 's that will occur and, by standard combinatorial arguments,

$$\mathbb{P}(Z_n(\lambda) = j) = \mathbb{P}(\text{exactly } j \text{ among } E_1, \dots, E_n \text{ occur}) = \sum_{i=j}^n (-1)^{i-j} \binom{i}{j} S_{i,n},$$

where

$$S_{0,n} = 1, \quad S_{i,n} = \sum_{1 \leq k_1 < \dots < k_i \leq n} \mathbb{P}(E_{k_1} \cap \dots \cap E_{k_i}), \quad i = 1, \dots, n.$$

Since the  $A$ 's are independent of the  $B$ 's, we have

$$\mathbb{P}(E_{k_1} \cap \dots \cap E_{k_i}) = \mathbb{P}(A_{k_1} \cap \dots \cap A_{k_i}) \mathbb{P}(B_{k_1} \cap \dots \cap B_{k_i}) = \lambda^i \frac{(n-i)!}{n!},$$

so that

$$S_{i,n} = \binom{n}{i} \lambda^i \frac{(n-i)!}{n!} = \frac{\lambda^i}{i!}, \quad i = 0, 1, \dots, n.$$

Therefore, the pmf of  $Z_n(\lambda)$  is given by

$$p_{n;\lambda}(j) := \mathbb{P}(Z_n(\lambda) = j) = \frac{1}{j!} \sum_{i=j}^n (-1)^{i-j} \frac{\lambda^i}{(i-j)!} = \frac{\lambda^j}{j!} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!}, \quad j = 0, 1, \dots, n. \quad (3.1)$$

The generalized matching distribution (3.1) has been introduced by [Niermann \(1999\)](#), who showed that  $p_{n;\lambda}$  is a proper pmf for all  $\lambda \in (0, 1]$ ; however, Niermann did not give a probabilistic interpretation to the pmf  $p_{n;\lambda}$ , and derived its properties analytically.

Since  $\lim_{n \rightarrow \infty} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!} = e^{-\lambda}$  for any fixed  $j$ , we see that  $p_{n;\lambda}$  converges pointwise to the pmf of  $Z(\lambda)$ , where  $Z(\lambda)$  is a Poisson rv with mean  $\lambda$ ,  $\text{Poi}(\lambda)$ . Interestingly enough, the Poisson approximation is extremelly accurate; numerical results are shown in [Niermann's \(1999\)](#) work. Also, Niermann proved that  $\mathbb{E} Z_n(\lambda) = \text{Var} Z_n(\lambda) = \lambda$  for all  $n \geq 2$  and  $\lambda \in (0, 1]$ . In fact, the following general result shows that the first  $n$  moments of  $Z_n(\lambda)$  and  $Z(\lambda)$  are identical, giving some light to the amazing accuracy of the Poisson approximation.

**Lemma 3.1.** *For any  $\lambda \in (0, 1]$ ,  $\mathbb{E}(Z_n(\lambda))_k = \lambda^k \mathbb{1}\{k \leq n\}$ ,  $k = 1, 2, \dots$ .*

*Proof.* For  $k > n$  the relation is obvious, since  $Z_n(\lambda) \in \mathcal{D}_n$ . For  $k = 1, 2, \dots, n-1$ ,

$$\mathbb{E}(Z_n(\lambda))_k = \sum_{j=k}^n \frac{\lambda^j}{(j-k)!} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!} = \lambda^k \sum_{r=0}^{n-k} \frac{\lambda^r}{r!} \sum_{i=0}^{(n-k)-r} \frac{(-\lambda)^i}{i!} = \lambda^k \sum_{r=0}^{n-k} p_{n-k;\lambda}(r),$$

and, since  $p_{n-k;\lambda}$  is a pmf supported on  $\{0, 1, \dots, n-k\}$ , we get the desired result. For  $k = n$ ,  $\mathbb{E}(Z_n(\lambda))_n = n! p_{n;\lambda}(n) = \lambda^n$ , completing the proof.  $\square$

**Corollary 3.1.** *For any  $\lambda \in (0, 1]$  and  $\alpha > 0$ ,*

$$\inf_{X \in \mathcal{D}_n} \{d_\alpha(X, Z(\lambda))\} = d_\alpha(Z_n(\lambda), Z(\lambda)).$$

Thus, for  $\lambda \in (0, 1]$ ,  $Z_n(\lambda)$  minimizes the factorial moment distance from  $Z(\lambda)$  over all rv's supported in a subset of  $\{0, 1, \dots, n\}$ . Using (2.5) it is easily verified that  $Z_n(\lambda)$  is unique. Moreover, it is worth pointing out that for  $\lambda > 1$ , we cannot find a random variable  $X \in \mathcal{D}_n$

such that  $\mathbb{E}(X)_k = \lambda^k \mathbb{1}\{k \leq n\}$  for all  $k$ . Indeed, since  $\mathcal{D}_n \subset \mathcal{X}(\infty) \subset \mathcal{X}(1)$ , assuming  $X \in \mathcal{D}_n$  and  $\mathbb{E}(X)_k = \lambda^k \mathbb{1}\{k \leq n\}$ , we get from (2.5) that

$$0 \leq \mathbb{P}(X = n-1) = \frac{\lambda^{n-1}(1-\lambda)}{(n-1)!},$$

which implies that  $\lambda \leq 1$ . Therefore, finding  $\inf_{X \in \mathcal{D}_n} \{d_\alpha(X, Z(\lambda))\}$  for  $\lambda > 1$  seems to be a rather difficult task.

We now evaluate some exact and asymptotic results for the factorial moment distance and the total variation distance between  $Z_n(\lambda)$  and  $Z(\lambda)$  when  $\lambda \in (0, 1]$ .

**Theorem 3.1.** Fix  $\alpha > 0$  and  $\lambda \in (0, 1]$  and let  $d_\alpha(n) := d_\alpha(Z_n(\lambda), Z(\lambda))$ . Then,

$$d_\alpha(n) = \frac{\alpha^n \lambda^{n+1}}{n!} \int_0^1 (1-y)^n e^{\alpha \lambda y} dy. \quad (3.2)$$

Moreover, the following double inequality holds:

$$\frac{\alpha^n \lambda^{n+1}}{(n+1)!} \left( 1 + \frac{\alpha \lambda}{n+2} + \frac{\alpha^2 \lambda^2}{(n+2)(n+3)} \right) < d_\alpha(n) < \frac{\alpha^n \lambda^{n+1}}{(n+1)!} \left( 1 + \frac{\alpha \lambda}{n+2} + \frac{\alpha^2 \lambda^2 e^{\alpha \lambda}}{(n+2)(n+3)} \right). \quad (3.3)$$

Hence, as  $n \rightarrow \infty$ ,

$$d_\alpha(n) \sim \frac{\alpha^n \lambda^{n+1}}{(n+1)!} \quad \text{and, more precisely,} \quad d_\alpha(n) = \frac{\alpha^n \lambda^{n+1}}{(n+1)!} \left( 1 + \frac{\alpha \lambda}{n+2} + o\left(\frac{1}{n}\right) \right). \quad (3.4)$$

*Proof.* From the definition of  $d_\alpha$  and in view of (1.5) and Lemma 3.1,

$$d_\alpha(n) = \frac{1}{\alpha} \sum_{k=n+1}^{\infty} \frac{(\alpha \lambda)^k}{k!} = \frac{1}{\alpha} \left( e^{\alpha \lambda} - \sum_{k=0}^n \frac{(\alpha \lambda)^k}{k!} \right) = \frac{1}{\alpha n!} \int_0^{\alpha \lambda} (\alpha \lambda - x)^n e^x dx,$$

and the substitution  $x = \alpha \lambda y$  leads to (3.2). Now (3.3) follows from the inequalities  $1 + \alpha \lambda y + \frac{1}{2} \alpha^2 \lambda^2 y^2 < e^{\alpha \lambda y} < 1 + \alpha \lambda y + \frac{1}{2} e^{\alpha \lambda} \alpha^2 \lambda^2 y^2$ ,  $0 < y < 1$ , while (3.4) is evident from (3.3).  $\square$

Theorems 2.1 and 3.1 give the next

**Corollary 3.2.** An upper bound of  $d_{\text{TV}}(Z_n, Z)$  is given by

$$d_{\text{TV}}(Z_n, Z) < \frac{2^n}{(n+1)!} \left( 1 + \frac{2}{n+2} + \frac{4e^2}{(n+2)(n+3)} \right) \sim \frac{2^n}{(n+1)!}. \quad (3.5)$$

The bound in (3.5) is of the correct order, and the same is true for the better result (1.1), given by DasGupta (1999, 2005). In contrast, the bound  $d_{\text{TV}}(Z_n, Z) \leq \frac{2^n}{n!}$ , given by Diaconis (1987), is not asymptotically optimal, because  $\frac{2^n}{(n+1)!} = o\left(\frac{2^n}{n!}\right)$ . Thus, it is of some interest to point out that the factorial distance  $d_2$  provides an optimal rate upper bound for the variational distance in the matching problem. The situation is similar for the generalized matching distribution, as the following result shows.



**Theorem 3.2.** For any  $\lambda \in (0, 1]$ , let  $d_{\text{tv}}(n) := d_{\text{tv}}(Z_n(\lambda), Z(\lambda))$  be the variational distance between  $Z_n(\lambda)$  and  $Z(\lambda)$ . Then,

$$d_{\text{tv}}(n) = \frac{\lambda^{n+1}}{2n!} \int_0^1 [y^n + (2-y)^n] e^{-\lambda y} dy. \quad (3.6)$$

Moreover, the following inequalities hold:

$$\begin{aligned} d_{\text{tv}}(n) &> \frac{2^n \lambda^{n+1}}{(n+1)!} \left( 1 - \frac{2\lambda}{n+2} \left( 1 - \frac{1}{2^{n+1}} \right) \right); \\ d_{\text{tv}}(n) &< \frac{2^n \lambda^{n+1}}{(n+1)!} \left( 1 - \frac{2\lambda}{n+2} \left( 1 - \frac{1}{2^{n+1}} \right) + \frac{4\lambda^2}{(n+2)(n+3)} \left( 1 - \frac{n+3}{2^{n+2}} \right) \right). \end{aligned} \quad (3.7)$$

Hence, as  $n \rightarrow \infty$ ,

$$d_{\text{tv}}(n) \sim \frac{2^n \lambda^{n+1}}{(n+1)!} \quad \text{and, more precisely,} \quad d_{\text{tv}}(n) = \frac{2^n \lambda^{n+1}}{(n+1)!} \left( 1 - \frac{2\lambda}{n+2} + o\left(\frac{1}{n}\right) \right). \quad (3.8)$$

*Proof.* Clearly, (3.8) is an immediate consequence of (3.7). Moreover, the inequalities (3.7) are obtained from (3.6) and the fact that  $1 - \lambda y < e^{-\lambda y} < 1 - \lambda y + \frac{1}{2} \lambda^2 y^2$ ,  $0 < y < 1$ . It remains to show (3.6). From (1.4) with  $p_1 = p_{n;\lambda}$  and  $p_2$  the pmf of  $\text{Poi}(\lambda)$ , we get

$$d_{\text{tv}}(n) = \sum_{j=0}^n \frac{\lambda^j}{j!} \left[ \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!} - e^{-\lambda} \right]^+ = \sum_{j=0}^n \frac{\lambda^j}{j!} \left[ \frac{(-1)^{n-j}}{(n-j)!} \int_0^\lambda (\lambda - x)^{n-j} e^{-x} dx \right]^+,$$

where the integral expansion is deduced by an application of (1.5) to the function  $f(\lambda) = e^{-\lambda}$ . Thus,

$$d_{\text{tv}}(n) = \sum_{k=0}^n \frac{\lambda^{n-k}}{(n-k)!} \left[ \frac{(-1)^k}{k!} \int_0^\lambda (\lambda - x)^k e^{-x} dx \right]^+ = \frac{1}{n!} \int_0^\lambda e^{-x} \left( \sum_{k \text{ even}} \binom{n}{k} (\lambda - x)^k \lambda^{n-k} \right) dx.$$

Since

$$\sum_{k \text{ even}} \binom{n}{k} (\lambda - x)^k \lambda^{n-k} = \frac{1}{2} [x^n + (2\lambda - x)^n],$$

we obtain

$$d_{\text{tv}}(n) = \frac{1}{2n!} \int_0^\lambda [x^n + (2\lambda - x)^n] e^{-x} dx,$$

and a final change of variables  $x = \lambda y$  yields (3.6).  $\square$

**Remark 3.1.** Although the factorial moment distance  $d_\alpha$  dominates the variational distance when  $\alpha \geq 2$ , the situation for  $\alpha \in (0, 2)$  is completely different. To see this, assume that for some  $\alpha \in (0, 2)$  and  $t \geq 0$  we can find a finite constant  $C = C(\alpha, t) > 0$  such that

$$d_{\text{tv}}(X_1, X_2) \leq C d_\alpha(X_1, X_2) \quad \text{for all } X_1, X_2 \in \mathcal{X}(t). \quad (3.9)$$

Obviously,  $Z$  and  $Z_n$ ,  $n = 1, 2, \dots$ , lie  $\mathcal{X}(\infty) \subset \mathcal{X}(t)$ . From Theorem 3.2 we know that  $\lim_{n \rightarrow \infty} \frac{(n+1)!}{2^n} d_{\text{tv}}(Z_n, Z) = 1$ . On the other hand, from (3.3) with  $\lambda = 1$ ,

$$d_{\alpha}(Z_n, Z) < \frac{\alpha^n}{(n+1)!} \left( 1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^{\alpha}}{(n+2)(n+3)} \right),$$

and, since  $|\alpha/2| < 1$ , this inequality contradicts (3.9):

$$1 = \lim_{n \rightarrow \infty} \frac{(n+1)!}{2^n} d_{\text{tv}}(Z_n, Z) \leq C \lim_{n \rightarrow \infty} \left( \left( \frac{\alpha}{2} \right)^n \left( 1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^{\alpha}}{(n+2)(n+3)} \right) \right) = 0.$$

## References

- BARBOUR, A.D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation*. Clarendon Press, Oxford University Press, New York.
- DASGUPTA, A. (1999). The matching problem with random decks and the Poisson approximation. *Technical report, Purdue University*.
- DASGUPTA, A. (2005). The matching, birthday and the strong birthday problem: a contemporary review. *J. Statist. Plann. Inference*, **130**, 377–389.
- DIACONIS, P. (1987). Application of the methods of moments in probability and statistics. *Proc. Sympos. Appl. Math.*, **37**, 125–142.
- NIERMANN, S. (1999). A generalization of the matching distribution. *Statistical Papers*, **40**(2), 233–238.
- WANG, Y.H. (1991). A compound Poisson convergence theorem. *Ann. Probab.*, **19**, 452–455.